

DOCUMENT RESUME

ED 186 968

CS 502 932

AUTHOR Boileau, Don M.; And Others
TITLE A Debate Judge Certification Test: Development and Operation on a State Wide Scale.
PUB DATE Nov 79
NOTE 15p.; Paper presented at the Annual Meeting of the Speech Communication Association (65th, San Antonio, TX, November 10-13, 1979).
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Certification; *Debate; Evaluation Criteria; *Evaluation Methods; *Judges; Persuasive Discourse; Qualifications; Secondary Education

ABSTRACT

In 1978-79 the Michigan Interscholastic Forensic Association adopted a three-part testing procedure to certify debate judges and make them eligible for officiating at district and state tournaments. The development and application of the procedure required a blending of debate principles and administration. Certification as a debate judge required attendance at an analysis session, passing a 50-item test on debate theory, and the successful completion of a sample ballot after viewing a videotaped debate. During the 1978-79 year of pilot-testing these procedures, 176 people took the certification test and an initial passing rate of 89.8% was achieved. During two more years of using the procedures, more evaluation data will be collected and revisions of the test forms will be made. (Author/RL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED186968

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A DEBATE JUDGE CERTIFICATION TEST: DEVELOPMENT
AND OPERATION ON A STATE WIDE SCALE

Written by:

Don M. Boileau,
Central Michigan University
Jon Fitzgerald, Executive Director,
Michigan Interscholastic Forensic Association
David Ling,
Central Michigan University
Dan P. Millar,
Central Michigan University

Abstract: In 1978-79 the Michigan Interscholastic Forensic Association adopted a 3-part testing procedure to certify debate judges who would be eligible to judge at the district and state tournaments. The development and application of the procedure required a blending of debate principles and administration. Certification requires attendance at an analysis session, passing a 50-item theory test, and successful completion of a ballot on a video-taped debate.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Don M. Boileau, Jon
Fitzgerald, David

Ling, Dan P. Millar
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Forensics Division: Competitive Program

Speech Communication Association

November, 1979

San Antonio, Texas

C5562932

A DEBATE JUDGE CERTIFICATION TEST: DEVELOPMENT
AND OPERATION ON A STATE WIDE SCALE

In the spring of 1978 the Executive Council of the Michigan Interscholastic Forensic Association (MIFA) decided that all judges in the 1978-1979 District and State Debate Tournament would have to meet minimum levels of competence as measured in a certification test.

Two growing pressures encourage the decision to implement a program of "judge competence"--accountability¹ and coach-judging.² Regarding the former, Ratliffe warns that "if the speech communication profession does not develop internal accountability we will continue as educators to be the victim of external pressure to be accountable on externally enforced terms."³ With the expense and visibility of high school debate, this area of communication education seemed particularly vulnerable to external pressures. With the trend in Michigan for coach-judging, a procedure was needed to validate the quality of persons making decisions affecting the students. The energies and aspirations of the student participating in competitive debate demanded the state association place qualified people in the position of "judge."

This paper provides an explanation of the development of the program and summarizes several operational decisions which were needed to implement the program. An overview gives the basic structure to the program which was implemented in 1978-79.

Overview

The adopted procedure for determining judge competency is based upon the assumption that a "qualified judge" should know the basic terminology found in the current topic, should know basic argumentation theory and the debate activity

rules, and should be able to write a ballot which reflected topic and theory knowledge in an educationally-sound manner.

The testing procedure occurs in three-stages. The first stage requires listening to an analysis of the current debate topic. Potential judges may attend workshops, conferences or special topic analysis sessions scheduled during the early weeks of the fall debate season. The purpose of the analysis is to familiarize potential judges with the areas for affirmative and negative cases and the specialized terminology of the proposition. Particular case positions are neither advocated nor condemned.

The second stage of the testing procedure requires that the potential judge score 80% on a 50-item written test. The test itself includes material from debate/argumentation theory, such as presumption, burden of proof, types of cases, tests of evidence and reasoning, as well as rules governing the debate procedure in Michigan. Each question is keyed to an argumentation/debate textbook or to the State rulebook. In this way, if a person fails to score 80%, they are informed of their errors and where they can find help in understanding the questions and answers. The purpose of the theory test is to insure that all judges have a minimum knowledge of debate and debating in order that high school students will not be judged by persons who literally do not know what they are doing.

The third stage requires that each potential judge view a video-tape of an actual high school debate and complete a ballot. The judge must fill out the entire ballot including the awarding of speaker points and ranks, writing individual comments to the four speakers and writing a rationale for the decision reported. What decision is made by the judge is less important than the basis for the decision and the individual remarks, the tone of the written comments and the specificity of the information provided. The underpinning value for these three general criteria is that an educationally-sound ballot should

contain sufficient information for the coach and debaters to reconstruct the debate and to understand WHY a particular vote was cast in a manner which is encouraging and conducive to student appreciation for the activity and skill development.

The certification procedure adopted by MIFA is modelled on the procedure used at Central Michigan University since 1973 to certify judges for the Department's Forensics Services--a service providing judges to the debate community surrounding the university. The development and upgrading of that certification procedure formed the basis for the program followed in the state. The present MIFA theory test grew out of work done at Central Michigan University to achieve minimum standards for CMU judges. Because each year a new group of graduate students with varying abilities and backgrounds arrived at CMU, the Forensics Services office needed to know in an objective fashion, which ones were qualified as judges. Also, a variety of faculty members from disciplines varying from math to broadcasting and community members varying from lawyers to housewives judged. Thus, in an attempt to develop some standardization, the three part certification program started in the fall of 1973.⁴

Written Test Construction

During this early experimentation it became apparent that the designators of "true" or "false" needed to be changed because test takers frequently perceived the question differently or could recall a specific incident which might change the context of the question. Thus, two major changes were made to tighten up the procedure. First, the terms were switched to "accurate" for "true" and "inaccurate" for "false." Thus, we would announce that accurate would mean "what was generally true of debate, but something for which there might be exceptions." Whereas "true" seemed to imply something which for most test takers was applied as always true. If a test taker recalled that in some debate back in high school

days, a judge allowed the opponents to answer from their seats, that person might answer false to the question: "Except for cross-examination the speaker will not be interrupted during the constructive or rebuttal speech." The word "accurate" allows the general rule to be applied for the correct answer.

The second major change was to allow the test taker to comment on the question in an attempt to develop the frame of reference for the answer. The person might then be able to show that the theory is understood, only that he/she is having difficulty with this particular wording of the question. Thus, in a question of inherency the test taker might write on the back of the sheet for that question, "If one is considering attitudinal inherency, the same restraint might reveal that the legislators will not change because of the power the opposition has to filibuster." Such a comment allows the test evaluator to decide whether to give credit for the question or not. In the MIFA test for one group of 53 people taking the test at two sites the average score of additional credit was 2. Although several people take advantage of the situation and write on almost twenty questions, most people write between 6-10 responses. Some candidates do not write any comments; this response may indicate confidence or lack of knowledge about the areas. Both explanations seem possible, as during the first year there were no written responses from two people with the highest scores (47 or higher) and one with a "not passing" score. This procedure does slow down the grading of the examination but is consistent with the nuances of debate.

During the fall of 1974 Ling and Boileau surveyed the debate coaches in Michigan at the Michigan Speech Association convention, to discover what areas of debate they felt should be covered in a test. Over forty topics were listed for the coaches to judge whether that topic was "essential for judging purposes," "helpful, but not essential," or "not necessary." From this list we learned from the 22 who responded that 12 essential areas existed. We still felt that certain

areas needed to be tested, so that we added seven more areas to give us our final areas for the test. (See table 1). For example, we originally did not think of the MIFA rules as an area of necessity. The comments of the coaches quickly established that many of the people coming from out of state would change decisions through some misunderstanding of the MIFA ballot or formats used in the state.

A validation of the written test was undertaken by a graduate student at the University of Washington for an independent study.⁵ He developed fifty questions from the short test at CMU and using the 12 areas specified by the Michigan coaches and the 7 areas which were added. This theory test was given to a group of the debate coaches at the 1977 National Forensic League national

TABLE 1
TOPIC AREAS FOR MIFA CERTIFICATION TEST

Comparative Advantage	Flow Sheet	Debate Procedure
Cross-Examination	Presumption	Tests of Evidence
Issues/Arguments/ Evidence	Inherency	Prima Facie
Negative Strategies	Refutation	Stock Issues
Burden of Proof	Reasoning	Delivery
Judging Styles	Critique	Ethics
MIFA Rules		

tournament in Seattle. He compared those results with a control group of people without debate, speech, or argumentation background. This control group came from people interested in debate by the fact that they had signed up for argumenta-

tion classes at the University of Washington. His results indicated 18 questions that distinguished between those with argumentation/debate backgrounds and those people without. Thus, we knew we had a core of questions which could not be answered by guessing. Many of the 32 other questions in this pilot seemed to have direction but did not reach the .05 level of confidence. The use of the "accurate/inaccurate" format allows people to guess, so wording becomes an important factor. It was from this list of 18 and the knowledge of the other 32 which did not distinguish of .05 that the final list of 50 questions were developed. The number of questions per topic area was decided by the test developed in terms of importance and the ability to ask questions of various areas within that topic.

The final pilot test of the theory examination came with the seven debate coaches on the MIFA council, which were the first group to take the final version of the test. Several questions were re-written to clarify wording since this group was asked to comment on the form of the question as well as answering it during the test.

Ballot Evaluation Procedure

The decision to require participants in the certification program to complete a written ballot of a video-taped debate was based on the assumption that a simple measure of knowledge of theory and practice (the objective examination) was not a sufficient demonstration of competence to judge a debate. In a competitive debate, judges are asked to do more than render a decision. They are expected to provide feedback of a nature that will allow the student to grow in his/her understanding of basic argumentation skills. This pedagogical function is most obvious when dealing with high school or beginning college debaters, but adheres generally in the concept of a critic/judge. Therefore, it was considered necessary to attempt to evaluate the quality of an individual's ballot-writing in order to ascertain two things. First, can the student/coach understand what

transpired during the debate from the perspective of the evaluator of that debate? Second, does the ballot provide comments that will allow the student to grow and develop basic skills of the reasoning process?

Ling and Boileau also asked the Michigan coaches to evaluate six different examples from debate ballots. From this feedback we learned that the specific comment which gave enough information for the coach who was not at the debate to understand the area being discussed was the type of comments desired. Short, general comments on skills without references to specific behavior or an argument did not have a lot of value for the coaches. These results were able to give us directions for our own evaluation of ballots.

In an attempt to achieve these goals the ballot evaluating process was established. Admittedly, this was the most subjective aspect of the certification process, since it involves such an obviously qualitative judgment. In an attempt to assure some consistency of evaluations, three evaluators reviewed each ballot.

Developing a video tape of the debate for obtaining a ballot to evaluate was a relatively easy procedure. Each year Central Michigan University sponsors a two-week summer debate workshop. The final round of that debate provided annually a new tape for testing purposes. The debate included both aspects of excellence as well as problems which people have with the topic. Since most debates are not championship level, a good debate with skilled, but not highly polished, people was desired. One year we substituted with a workshop debate from a one-day workshop in September. A close debate which allows reasonable justification of either a negative or affirmative decision functions best.

Each evaluator is asked to judge the ballot on the following three criteria: 1) quality, 2) specificity, and 3) tone. Each of these criteria is described below.

Quality of the Comments: Do the bases of decisions and the comments to the individual debaters reflect what occurred during the debate? (Were critical

errors not mentioned? Were major issues ignored? etc.) The question is not "Did the judge vote correctly?" Indeed, debates selected for this phase of certification were intended to be close enough for competent judges to disagree about the outcome. One of the tapes used in 1978-79, for example, has had a 30-70 split of affirmative and negative decisions. The concern is that the comments are directed at the major issues in such a way that someone who had not observed the debate would be able to understand how the judge arrived at a given decision. The intent is to avoid ballots that are vague or reflect the issues that the judge would have preferred to see ("debating the debaters").

Consider the following basis of decision:

The affirmative case withstood negative attack; negative could not establish that the status quo can solve problems inherent to the system--Plan attacks were either answered or left without impact.

This basis of decision could describe almost any debate in which the affirmative was victorious and tells neither the victor nor the vanquished what they did correctly or incorrectly. Contrast such a ballot with the following from one of the ballots:

Negative effectively reduced the impact of the finite nature of fossil fuels and combined with a diversified use of energy sources decreases significantly aff. philosophy of dependence on finite fossil fuels. Through use of updated evidence, negative establishes decreasing trend of harmful effects of fossil fuels in terms of technology and legislation. Affirmative does show a threat of economic decline in growth. However, negative minimizes by showing very long term nature.

Whether one agrees or disagrees with the evaluation, a debater of his/her coach knows what issues emerged in the debate and how they were evaluated. Now the coach has a basis for instructing the affirmative on what needs to be adjusted and for teaching the negative what issues need to be emphasized in similar situations.

Specificity and constructiveness of comments: Is information specific enough

so that the students and their coaches can utilize the ballot as an aid in the student's growth? Most high school debate ballots now include an area for judges to make comments specifically to each of the debaters. Here the judge is asked to aid directly in the education of the students. But such comments are only helpful if they are specific and constructive, the kind of comments that can be readily utilized. Notice in the following comments to a second negative how the critic identifies strengths and weaknesses and also indicates directions the student might pursue.

Plan-meet-advantage arguments are well suited to plan; impact statements could be more forceful if you delineated more specifically the impact of your indictments. Disadvantage #1; isolating impact on numbers of unemployed would increase your impact because you could delineate more specifically who was harmed. Disadvantage #2; link in D₂ is very weak. The impact on world trade being upset and destruction of West would also strengthen argument.

While such comments provide a helpful aid to the student's development, comments like "Very good job." "You have a pleasant voice and you are well organized" (when these are the only comments on the ballots) provide little that can be used to help the student progress.

Tone: Are the comments designed to encourage the student to grow in the activity? This third criteria is the least precise and most difficult to exemplify. Here the concern is that the judge recognize and be sensitive to the fact that they are addressing young high school students. Caustic comments like, "This constructive speech seemed to be a waste of negative team's time" or "Did you ever consider band?" are among blatant instances of judges potentially destroying a student's future participation in the activity. The importance of such a criterion is highlighted by the fact that judges often are unaware of the tone of their written communications. Developing a written style that criticizes without derogating is an important skill. For that reason, tone is included in the criteria for certification.

After an evaluator reads the ballot, he writes a short critique of the ballot along with a pass/fail judgment. If two of the three evaluators judge the ballot to be unacceptable, the individual is required to repeat this phase of certification at a later date. The possibility exists that the next ballot will be written "to the evaluators," and that the individual will thereafter revert to the past unacceptable style of ballot-writing. However, it is the hope that the individual desires to improve his/her ballot-writing skills and will make a sincere effort to adjust his/her future performance. Should such not be the case, there is an additional check on such reversions. If a judge submits a ballot that a coach believes fails to meet the minimum standards, the coach may submit a complaint along with the ballot to the office of the Executive Director of the Michigan Interscholastic Forensic Association. If repeated complaints are received, the Executive Council of MIFA reserves the right to withdraw in individual's certification. None were received this year.

FIRST YEAR OF OPERATION:

The MIFA council discussed the possibility of using this procedure at three different meetings. Before a decision was made, a program at the spring conference of the Michigan Speech Association on the proposal allowed the test developers to meet with interested high school debate coaches. The precise motion and the rationale as well as a sample theory test were reviewed in the meeting. At the final meeting before adoption, as mentioned earlier, the high school coaches went through the exact testing procedures, so they could experience the test during examination conditions.

After adoption, an announcement was mailed in May to all the coaches. Sites for testing were established to give geographical and program balance. The first site was a large tournament at the end of the year. Two different meetings of the Michigan Speech Association were selected to give other opportunities.

The University of Michigan/Michigan Speech Association summer conference was a time and location of the largest summer debate institute, so several coaches would be there in addition to the many coaches serving on the staff. The annual fall meeting of the Michigan Debate Coaches conference, which coincides with the state wide Student Congress, was a large test group. On site tests were given at one league meeting area on the west side of the state and another in the Upper Peninsula. Because the number of coaches that were expected to take the test was not reached by the middle of October, we also gave it at the state novice tournament in December and an invitational tournament late in November in the Detroit area. This latter site allowed many of the non-coaches who judge during the season to take the test before districts. (These non-coaches are mainly undergraduate students, law students, graduate students from other areas with debate backgrounds, and lawyers.)

To help in the diffusion process we put an explanation in the Michigan Bulletin of Secondary School Principals. Several principals were then willing to pick up the \$5 test fee, and/or encourage their coach to take the test. This announcement was part of the accountability development of the program.

To standardize the giving of the test, a set of written instructions were developed. Since five different people gave the test during the season, these instructions were valuable in standardizing the format. Also, a check list of equipment for the VTR were available. The materials needed were: 1) the written test copies, 2) the IBM scoring sheets, 3) the registration form, 4) receipts, 5) the ballots, 6) the video tape, and 7) the instructions. For most viewings we used two TV monitors so that people would be able to see and hear well.

Test takers were asked to write their question qualification on the back side of the IBM sheets. Flow sheets were collected afterwards to help with test security. On two occasions we were able to review the test taker's flow sheet and determine that the person not passing the ballot part of the test had taken a

thorough flow sheet and only needed to take the ballot part of the test again. Thus, in two cases, the flow sheets functioned to indicate that the problem was not the flow sheet providing enough material for the ballot, while in a third case the inadequate flow sheet indicated that the person needed more exposure to debate before judging.

Analysis sessions were allowed in three distinct fashions. First, a person working and/or attending a summer debate institute could qualify. Second, a person attending an analysis session at the debate coaches meeting, the Michigan Speech Association program on this year's topic, or a person attending an analysis session at one of the one-day workshops in the fall that several of the universities sponsor. Third, a person could purchase for \$3.00 a tape recording of an analysis session.

During the year, 176 people took the test. The initial passing rate was 89.8 percent. Six people passed the test on the second attempt (5 on the ballot--1 on the theory). The state tournament used 79 certified judges. Ten of the coaches with teams in the state tournament had to hire certified judges as they were not certified.

MIFA has undertaken a three-year commitment to this program. Thus, final evaluation data has not been completed. Item analysis of the theory test has been undertaken and will lead to a revision of the test forms which were used.

Foot Notes

¹Jane Work, "Professional Accountability in Speech Communication: Promise, Plaything or Primrose Path," Communication Education, 25, (January 1976), 68-74.

²Don M. Boileau and David A. Ling, "Judge Certification," MIFA Newsletter, Fall 1974, 6-8.

³As quoted in Work, p. 69.

⁴A short 16 question test was used during the first year. Both multiple choice and true/false questions were used; a 90 percent score was required. Each person taking the test went over the results with one of the members of the steering committee (all four were or had been debate coaches--3 of whom have had teams at the national tournament). This short test was revised each of the first four years it was tried. The revisions were based on the feedback from the conferences with the test takers.

⁵Val Flake, "Comparison of Experienced and Non-Experienced Debate Judges on a 50-item Theory Test," (unpublished, University of Washington independent study paper, 1977).